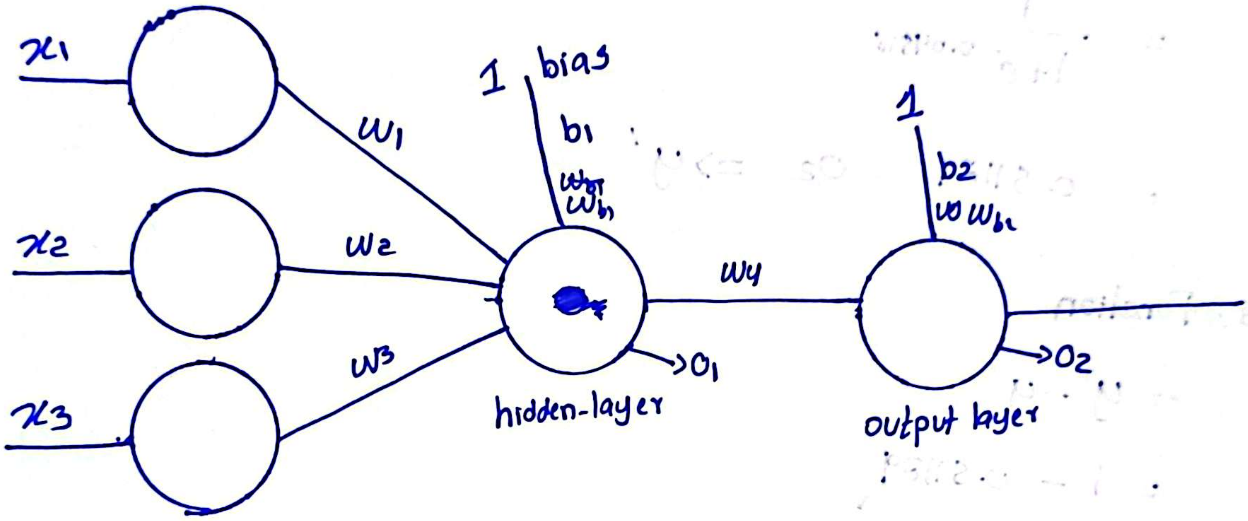
Artificial Neural Network

x_1 IQ	x_2 Study-hours	x_3 Play-hours	output
95	4	4	1
100	5	2	1
95	2	7	0



Forward Propagation

Lets say: $w_1 = 0.01, w_2 = 0.02, w_3 = 0.03, w_{b_1} = 0.001, w_4 = 0.02,$
$w_{b_2} = 0.001$

Step:1
Hidden layer

$$z = \sum_{i=1}^{n} w_i x_i + b$$

$$z = (95 \times 0.01) + (4 \times 0.02) + (4 \times 0.03) + (1 \times 0.001)$$

$$= 1.151$$

Step:2: Activation Function

Let use sigmoid

$$f(z) = \frac{1}{1+e^{-z}}$$

$$f(z) = \frac{1}{1+e^{-1.151}} = 0.759 = O_1$$

Hidden layer 2

Step : 1

$$z = O_1 \times w + b$$

$$= (0.759 \times 0.02) + (1 \times 0.03)$$

$$= 0.04518$$

Step 2

$$= \frac{1}{1+e^{-0.04518}}$$

$$= 0.51129 = O_2 \Rightarrow y'$$

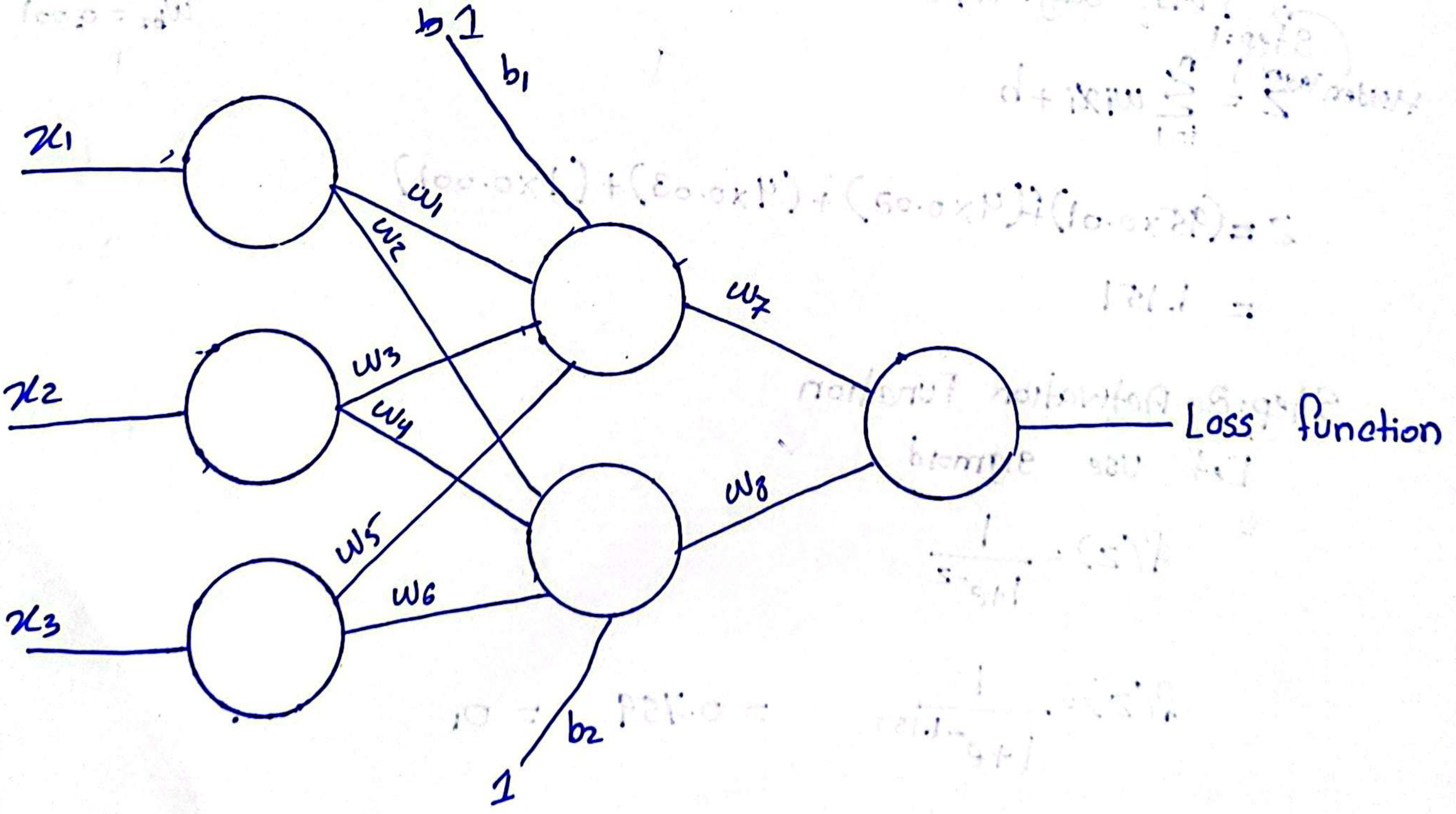
Loss Function

$$= y - y'$$

$$= 1 - 0.51129$$

$$\approx 0.49$$

$\longrightarrow$ Back Propagation and Weight Updatation



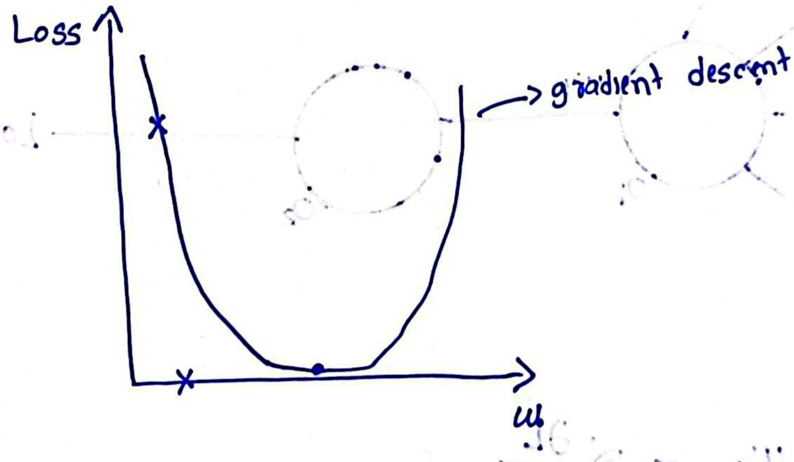
$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \\ w_5 & w_6 \end{bmatrix}$$

Weight Updation Formula

$$w_{7_{new}} = w_{7_{old}} - \eta \underbrace{\left[\frac{\partial L}{\partial w_{7_{old}}}\right]}_{\text{slope}}$$

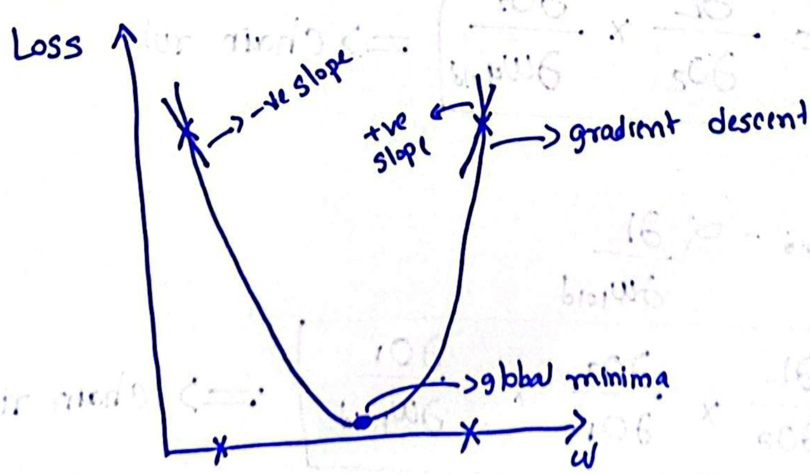
$$w_{3_{new}} = w_{3_{old}} - \underbrace{\eta}_{\text{learning rate}} \frac{\partial L}{\partial w_{3_{old}}}$$

$$\boxed{w_{new} = w_{old} - \eta \frac{\partial L}{\partial old}} \Longrightarrow \text{weight updation formula}$$



Optimizers

Optimizer role is to reduce the loss value



-ve slope

$$w_{new} = w_{old} - \eta(-ve)$$

$$= w_{old} + \eta(+ve)$$

$$w_{new} \geq w_{old}$$

+ve slope

$$w_{new} = w_{old} - \eta(+ve)$$

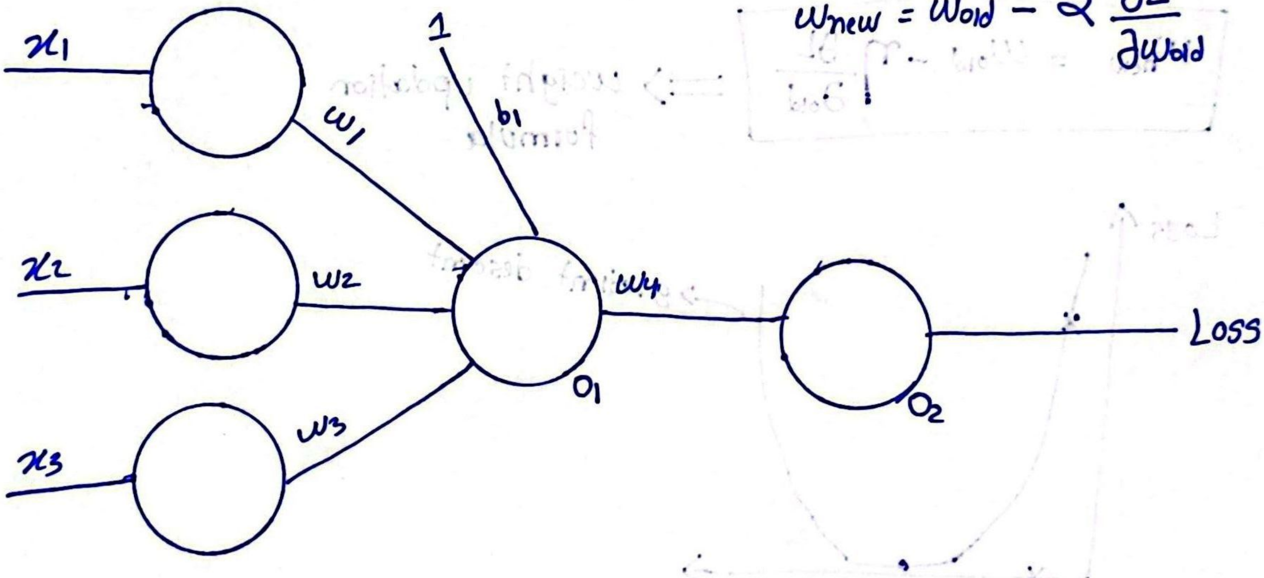
$$= w_{old} - \eta(+ve)$$

$$w_{new} < w_{old}$$

when w reaches global minima

$$w_{new} = w_{old}$$

⟶ Chain Rule of Derivatives.



$$w_{new} = w_{old} - \underbrace{\alpha}_{\text{learning rate}} \frac{\partial L}{\partial w_{old}}$$

$$w_{4_{new}} = w_{4_{old}} - \alpha \frac{\partial L}{\partial w_{4old}}$$

$$\boxed{\frac{\partial L}{\partial w_{4old}} = \frac{\partial L}{\partial O_2} \times \frac{\partial O_2}{\partial w_{4old}}} \Rightarrow \text{chain rule}$$

$$w_{1new} = w_{1old} - \alpha \frac{\partial L}{\partial w_{1old}}$$

$$\boxed{\frac{\partial L}{\partial w_{1old}} = \frac{\partial L}{\partial O_2} \times \frac{\partial O_2}{\partial O_1} \times \frac{\partial O_1}{\partial w_{1old}}} \Rightarrow \text{chain rule}$$